

Boundaries of Hypertrees, and Hamiltonian Cycles in Simplicial Complexes

Rogers Mathew¹, Ilan Newman², Yuri Rabinovich², and Deepak Rajendraprasad²

¹ Department of Computer Science, Indian Institute of Technology Kharagpur, India.

² Department of Computer Science, University of Haifa, Israel.

December 8, 2015

Abstract

A d -hypertree on $[n]$ is a maximal acyclic d -dimensional simplicial complex with full $(d-1)$ -skeleton on the vertex set $[n]$. Alternatively, in the language of algebraic topology, it is a minimal d -dimensional simplicial complex T (assuming full $(d-1)$ -skeleton) such that $H_{d-1}(T; \mathbb{F}) = 0$.

The d -hypertrees are a basic object in combinatorial theory of simplicial complexes. They have been studied; and yet, many of their structural aspects remain poorly understood.

In this paper we study the boundaries $\partial_d T$ of d -hypertrees, and the fundamental d -cycles defined by them. Our findings include:

- A full characterization of $\partial_d T$ over \mathbb{F}_2 for $d \leq 2$, and some partial results for $d \geq 3$.
- Lower bounds on the maximum size of a largest simple d -cycle on $[n]$. In particular, for $d = 2$, we construct a *Hamiltonian d -cycle* H on $[n]$, i.e., a simple d -cycle of size $\binom{n-1}{d} + 1$. For $d \geq 3$, we construct a simple d -cycle of size $\binom{n-1}{d} - O(n^{d-2})$.
- Observing that the maximum of the expected distance between two vertices chosen uniformly at random in a tree (1-hypertree) on $[n]$ is at most $\sim n/3$, attained on Hamiltonian paths, we ask a similar question about d -hypertrees. *How large can be the average size of a fundamental cycle of a d -hypertree T (i.e., the expected size of the dependency created by adding a d -simplex on $[n]$, chosen uniformly at random, to T)?* For every $d \in \mathbb{N}$, we construct an infinite family of d -hypertrees $\{T\}$ with the average size of a fundamental cycle at least $c_d |T| = c_d \binom{n-1}{d}$, where c_d is a constant depending on the dimension d alone.

1 Introduction

Graph Theory, and in particular its portion dealing with connectivity-related notions such as cycles, trees and cuts, and their packing, covering, polytopes, etc., plays a fundamental role in CS, and is at the core of Combinatorial Optimization and the theory of Algorithms. While graphs are well-suited for modeling systems with pairwise interactions, modeling of more complex multiway interactions is required in fields like Game Theory, Distributed Systems, Image Processing, etc. This calls for a robust higher-dimensional generalization of graphs. Hypergraphs provide a partial answer, however, it will be noted that the graph-theoretic connectivity notions do not even have a widely-agreed formulation in this context, much less a coherent theory.

In contrast, simplicial complexes possess a richer structure than hypergraphs, and do allow a natural and meaningful generalization of trees, cycles, etc. In fact, both simplicial *homology* and *homotopy* theories can be viewed as high-dimensional connectivity theories. This paper is dedicated to the study of the combinatorial structure of these generalizations, namely d -hypertrees and d -cycles.

On simplicial complexes and their applications. Simplicial complexes have been introduced by Poincaré at the turn of the 20'th century as a tool for creating a rigorous foundation for what has later evolved into the homology theory for smooth manifolds. The introduction of clear combinatorial notions had immediately led to flourishing of Combinatorial Topology. Although later the simplicial

methods were gradually replaced by a powerful algebraic apparatus, simplicial complexes have retained their importance in this branch of mathematics up to this day. Gradually, simplicial complexes along with the associated topological methods spread through virtually all branches of Discrete Math. and Computer Science¹.

In recent years, the study of simplicial complexes has continued at an ever accelerating rate. A series of brilliant papers dedicated to the study of hard topological invariants like embedding dimension and homotopy groups, see e.g., [22, 6, 21], has significantly promoted our understanding in this direction. Spectacular advances in the study of threshold phenomena in random complexes, see e.g. [16, 4, 3, 2, 1, 18], have virtually shaped a new important topological-combinatorial area out of thin air. Another important similar development had to do with constructing expanders of various types, see e.g., [11, 9, 15, 20]. In practical CS, topological methods for data analysis have become exceedingly popular. For an overview of this field see, e.g., the summaries [7, 10]. In addition to classical methods of reconstructing geometrical data from properly produced samples (see, e.g., [8]), new methods based on *persistent homology* have emerged, and become widespread.

To sum up, the theory of simplicial complexes provides valuable tools for a large variety of theoretical and practical areas. Situated at the meeting place of Geometry and Combinatorics, it provides a language understandable to both, and serves as an important vehicle for exchange of ideas and results between the two areas.

Our contribution. In view of the above, it is surprising that the combinatorial structure of simplicial complexes is rather poorly understood at the moment, and even the simplest questions about the structure of the basic objects such as 2-hypertrees, 2-cycles and 2-cocycles, run into unknown territory.

A word about d -hypertrees is due. While d -cycles and d -cocycles are standard and very basic notions of Combinatorial/Algebraic Topology requiring no introduction, the notions of hypertree, although standard, is less common. A d -hypertree T is a maximal d -acyclic subcomplex of the complete d -dimensional complex K_n^d on the vertex set $[n]$. This definition implies, in particular, that T has a full $(d-1)$ -skeleton. Equivalently, a d -hypertree is a minimal subcomplex $K_n^{d-1} \subset T \subseteq K_n^d$ such that $\tilde{H}_{d-1}(T, \mathbb{F}) = 0$ with respect to the underlying field \mathbb{F} . Hypertrees were introduced and first studied by Kalai in [14], and were subsequently studied, e.g., in [16, 24, 17, 25, 26].

This paper (and its companion [19]) is close in spirit to the afore-mentioned research of the threshold probabilities of various properties of random d -complexes. We study a number of extremal-type problems pertaining to the combinatorial/topological structure of d -hypertrees and d -cycles.

Our first goal is the characterization of the boundaries of d -hypertrees. Although at first glance it may sound a rather exotic question, its solution provides an approach to solving mainstream problems about d -hypertrees and d -cycles. For $d = 2$ we obtain a complete characterization; for larger d 's an approximate answer is provided.

Next, we study simple d -cycles of the maximum size over the vertex set $[n]$. Clearly, the size of such a cycle cannot exceed the size of a d -hypertree by more than 1. Call such d -cycles *Hamiltonian*. As in graphs, and for the same reason, all the d -hypertrees over $[n]$ have the same size $\binom{n-1}{d}$. Call also a d -hypertree *Hamiltonian* if it is obtainable from a Hamiltonian d -cycle by removing one of its d -simplices. Do such d -cycles and d -hypertrees exist? We show that for $d = 2$ the answer is positive if and only if $n \equiv 0$ or $3 \pmod{4}$. For larger d 's we can only show that there exist simple d -cycles of size at least $\binom{n-1}{d} - O(n^{d-2})$.

Hamiltonian d -hypertrees as defined above are just one possible generalization of Hamiltonian paths in graphs. For another, metric generalization, observe that Hamiltonians paths are the extremal (connected) graphs with respect to the expected distance between a random pair of vertices u, v , attaining value $\sim n/3$. Interpreting the distance between the vertices u and v in a tree T as the size (minus 1) of the fundamental cycle obtained by adding the edge (v, u) to it, one naturally arrives to the following problem about d -hypertrees. How large can be the average size of an fundamental cycle of a

¹See, e.g., the excellent book of Matoušek [23], and also [5], for some beautiful examples of uses of topological methods in Graph Theory and vice versa. The most famous applications of simplicial complexes in Theoretical CS include the study of evasiveness of nontrivial monotone graph properties of Kahn et al. [13], and the consensus problem in the area of asynchronous concurrent computation [12, 27]

d -tree T (i.e., the dependency created by adding a random uniform d -simplex on $[n]$ to T)? For every $d \in \mathbb{N}$, we explicitly construct a d -hypertree T attaining value $\geq c_d \binom{n-1}{d}$ for some c_d depending only on d . I.e., a constant fraction of the fundamental cycles with respect to T are of size commensurate with T .

We work mostly with an interesting subclass of collapsible d -hypertrees, systematically employing a technique which we call the conical extension. Both this class of d -hypertrees, and the technique used, are potentially useful for other hypertree-related constructions.

1.1 Simplicial complexes

Some standard notation. We use \mathbb{F}_2 to denote the two element finite field and \mathbb{Q} to denote the field of rational numbers. The notation $[n]$ is a shorthand for the set $\{1, \dots, n\}$. If A and B are sets, then $A \oplus B$ denotes their symmetric difference and if A and B are vectors over \mathbb{F}_2 , then it denotes their vector sum.

A d -dimensional simplex (d -simplex or d -face for short) is a set with $d + 1$ elements. A simplicial complex X is a collection of simplices that is closed under containment. That is, if $A \in X$ then every subset of A also belongs to X . The union of all the simplices in X is called the *vertex-set* $V(X)$ of X . In this article, we will always assume that $V(X)$ is finite and identify it with $[n]$, where $n = |V(X)|$. The *dimension* of a simplicial complex X is the largest dimension among all the simplices in X . Further, X is called *pure* if all its maximal faces are of the same dimension. The set of i -dimensional simplices of X is denoted by $X^{(i)}$. The *complete n -vertex d -dimensional simplicial complex* $K_n^d = \{\sigma \subset [n] : |\sigma| \leq d + 1\}$ contains all the simplices in $[n]$ with dimension at most d .

Chains. Given a field \mathbb{F} and a simplicial complex X , an \mathbb{F} -weighted formal sum Z of the d -faces in X is called a d -chain over \mathbb{F} , i.e., $Z = \sum_{\sigma \in X^{(d)}} c_\sigma \sigma$, where $c_\sigma \in \mathbb{F}, \forall \sigma \in X^{(d)}$. The set $\text{supp}(Z)$ of d -simplices with a non-zero weight in Z is called its *support* and $|\text{supp}(Z)|$ is called the *size* $|Z|$ of Z . The vertex-set of X is also referred to as the vertex-set of Z . The collection of all d -chains of K_n^d form a vector space \mathcal{C}_d over \mathbb{F} with dimension $\binom{n}{d+1}$.

The boundary operator. The *boundary* $\partial_d \sigma$ of a d -simplex $\sigma = \{v_0, \dots, v_d\}$, with $v_0 < \dots < v_d$, is the $(d - 1)$ -chain $\sum_{i=0}^d (-1)^i (\sigma \setminus \{v_i\})$. The signs disappear when one works over \mathbb{F}_2 . The linear extension of the same to the whole of \mathcal{C}_d is the *boundary operator* $\partial_d : \mathcal{C}_d \rightarrow \mathcal{C}_{d-1}$. It is a well known and easily verifiable fact that $\partial_{d-1} \partial_d = 0$. The subscript of the operator may be omitted when unambiguous.

Cycles, forests and trees. A d -chain Z is called a d -cycle if $\partial Z = 0$. We refer to $0 \in \mathcal{C}_d$ as the *trivial d -cycle*. Further, if there is no nontrivial cycle supported on a proper subset of $\text{supp}(Z)$, then Z is called *simple*. A set F of d -faces is called a d -forest (or *acyclic*) over a field \mathbb{F} if F does not contain the support of any nontrivial d -cycle over \mathbb{F} . Further, F is called a d -hypertree or a d -tree on $[n]$ if it is a maximal d -forest in K_n^d under inclusion. It is easy to see that for a d -forest F , the set of $(d - 1)$ -chains $\{\partial \sigma : \sigma \in F\}$ is linearly independent in \mathcal{C}_{d-1} and when F is a d -tree this set becomes a basis for the cycle-space in K_n^{d-1} , that is the kernel of ∂_{d-1} . Hence all d -trees on $[n]$ have size $\binom{n-1}{d}$.

Duals and cuts. The *dual* of a d -chain $Z = \sum_{\sigma \in \text{supp}(Z)} c_\sigma \sigma$ over $[n]$ is the $(n - d - 2)$ -chain $Z^* = \sum_{\sigma \in \text{supp}(Z)} c_\sigma \bar{\sigma}$, where $\bar{\sigma} = [n] \setminus \sigma$. A *cut* is the dual of a simple cycle. Over \mathbb{F}_2 , the dual of a set of d -simplices Z over $[n]$ is the set Z^* of $(n - 2 - d)$ -simplices $\{[n] \setminus \sigma : \sigma \in Z\}$.

Degree. The *degree* of an i -face σ in a simplicial complex X is the number of $(i + 1)$ -faces in X which contains σ . The *maximum (resp., minimum) degree* of a d -dimensional complex X is the largest (resp., smallest) degree among all its $(d - 1)$ -faces.

Collapsibility. In a d -dimensional simplicial complex X , a $(d - 1)$ -face τ is called *exposed* if its degree is 1, that is, it belongs to exactly one d -face σ of X . An elementary d -collapse on τ consists of the removal of σ and τ from X . We say that X is *d-collapsible* if every d -face of X can be removed by a sequence of elementary d -collapses. It is easy to see that if X is d -collapsible, then $X^{(d)}$ is a d -forest over any field.

Hamiltonian cycles and trees. A d -cycle Z on $[n]$ is called *Hamiltonian* if it is simple and has size $\binom{n-1}{d} + 1$. Notice that this is the largest possible size of a simple d -cycle on $[n]$. Removing any d -face from the support of Z results in a d -tree. Such d -trees, i.e., d -trees contained in the support of a Hamiltonian d -cycle, are called *Hamiltonian d-trees*.

Fillings and filling-volume. For any $(d - 1)$ -cycle Z and any d -tree T , both on $[n]$, there exists a unique d -chain F supported on T such that $\partial F = Z$. This d -chain F is called the *filling* of Z in T and is denoted by $\text{Fill}(Z, T)$. When Z above is the boundary $\partial\sigma$ of some d -simplex $\sigma \in K_n^d$, we call $\text{Fill}(\partial\sigma, T)$ as the filling of σ in T and denote it by just $\text{Fill}(\sigma, T)$. The *average filling-volume* $\mu(T)$ of T is $\binom{n}{d+1}^{-1} \sum_{\sigma \in K_n^d} |\text{Fill}(\sigma, T)|$.

Abuses of notation (Important). The first type of abuse is to blur the distinction between a pure d -dimensional complex X and its set of d -faces $X^{(d)}$. For example, we might say that X is a forest (resp., a tree) to mean that $X^{(d)}$ is a d -forest (resp., a d -tree). In the other direction, if Y is a collection of d -faces, by collapsibility (resp., maximum degree, minimum degree) of Y we mean the collapsibility (resp., maximum degree, minimum degree) of the simplicial complex obtained by taking the subset-closure of Y . The second type of abuse is to blur the distinction between a d -chain and its support when working over \mathbb{F}_2 . We will say more about this in Section 1.2.

1.2 Preliminaries

Working over the field \mathbb{F}_2

A d -chain over \mathbb{F}_2 can be identified bijectively with its support. Addition of two chains will correspond to the symmetric difference of their supports. Boundary of a collection of d -simplices X is understood as the boundary of the unique d -chain over \mathbb{F}_2 whose support is X . It may be helpful to note that, in this case, the boundary of a d -simplex is just the collection of $(d - 1)$ -simplices contained in it and thus ∂X is the collection of $(d - 1)$ -faces with an odd degree in X . Henceforth in this article, we will work only over \mathbb{F}_2 , although our results extend to \mathbb{Q} (cf. Section 5).

Basic operators and related notions

Besides the boundary operator ∂ discussed above, additional standard and less standard operators will be used in this paper.

Link. The *link* operator maps a simplicial complex X to the neighborhood of a vertex v :

$$\text{Link}_v(X) = \{\sigma \setminus \{v\} \mid \sigma \in X, v \in \sigma\}.$$

With a slight abuse of notation, we shall treat $\text{Link}_v(X^{(d)})$ as an operator mapping a collection $X^{(d)}$ of d -simplices over $[n]$, to a collection of $(d - 1)$ -simplices over $[n] \setminus \{v\}$. Observe that a link of a d -cycle Z is a $(d - 1)$ -cycle, since in this case $\text{Link}_v(Z) = \partial\{\sigma \mid v \in \sigma \in Z\}$, and $\partial\partial = 0$.

Cone. The *cone* operator is the right inverse of the link operator; it maps a collection $X^{(d)}$ of d -simplices over $[n]$, to a collection of $(d + 1)$ -simplices over $[n] \cup \{v\}$, where $v \notin [n]$:

$$\text{Cone}_v(X^{(d)}) = \left\{ \sigma \cup \{v\} \mid \sigma \in X^{(d)} \right\}.$$

It is easy to verify that the boundary of $\text{Cone}_v(Y)$ is given by

$$\partial \text{Cone}_v(Y) = Y \oplus \text{Cone}_v(\partial Y). \tag{1}$$

Conical extension. Let $X^{(d)}$ and $Y^{(d-1)}$ be collections of d - and $(d-1)$ -simplices, respectively, over $[n]$, and let $v \notin [n]$. We define the *conical extension* $\text{Ext}_v(X^{(d)}, Y^{(d-1)})$, a collection of d -simplices on $[n] \cup \{v\}$, by

$$\text{Ext}_v(X^{(d)}, Y^{(d-1)}) = X^{(d)} \oplus \text{Cone}_v(Y^{(d-1)}).$$

(The two summands are in fact disjoint.) It is easy to verify using (1), that if both $X^{(d)}$ and $Y^{(d-1)}$ are acyclic, then so is $\text{Ext}_v(X^{(d)}, Y^{(d-1)})$. Similarly, if both $X^{(d)}$ and $Y^{(d-1)}$ are collapsible, then so is $\text{Ext}_v(X^{(d)}, Y^{(d-1)})$.

Moreover, assume that $X^{(d)}$ and $Y^{(d-1)}$ are, respectively, d - and $(d-1)$ -hypertrees on $[n]$. Keeping in mind that an acyclic collection of k -simplices on $[n]$ is a k -hypertree if and only if its size is $\binom{n-1}{k}$, we conclude that $\text{Ext}_v(X^{(d)}, Y^{(d-1)})$ is a d -hypertree on $[n] \cup \{v\}$, since it is acyclic, and its size is $\binom{n-1}{d} + \binom{n-1}{d-1} = \binom{n}{d}$.

This leads to the following purely combinatorial definition-theorem, fundamental for this paper:

Definition 1.1 (Nice trees). Let $\{T_k\}_{k=d+1}^{n-1}$ be a sequence of $(d-1)$ -hypertrees, where T_k has an underlying vertex set $[k]$, respectively. Let $H_{d+1} = \{1, \dots, d+1\}$ (a single d -simplex), and define recursively, for $k = d+1, \dots, n-1$,

$$H_{k+1} = \text{Ext}_{k+1}(H_k, T_k).$$

Then, each H_k is a d -hypertree on $[k]$. Call such hypertrees *nice*.

They form a subfamily of collapsible d -hypertrees. For $d = 1$, all trees are nice.

Filling and conical extension. The last operator to be discussed in this section is $\text{Fill}(*, T)$, a bijection between the $(d-1)$ -cycles on $[n]$ and subsets of d -simplices of a fixed d -tree T . Let $Z \in \mathcal{Z}_{d-1}$ be a $(d-1)$ -cycle on $[n]$. Then, by the definition of d -hypertree, there exists a unique subset F of d -simplices in T , such that $\partial F = Z$. We define $\text{Fill}(Z, T) = F$. Due to acyclicity of T , this is indeed a bijection. Observe that the operator $\text{Fill}(*, T)$ is linear: $\text{Fill}(Z_1 \oplus Z_2, T) = \text{Fill}(Z_1, T) \oplus \text{Fill}(Z_2, T)$. Finally, to put things in a graph-theoretic perspective, observe that for a d -simplex σ on $[n]$, $\text{Fill}(\partial\sigma, T) \oplus \sigma$ is an analogue of what is called in graph theory *a fundamental cycle of σ with respect to T* .

The following claim about fillings of conical extensions will be used in Section 4.

Claim 1.2. Let T_{n-1}^d and T_{n-1}^{d-1} be, respectively, a d -hypertree and a $(d-1)$ -hypertree on $[n-1]$. They define, by means of conical extension, a new d -hypertree T_n^d on $[n]$: $T_n^d = \text{Ext}_n(T_{n-1}^d, T_{n-1}^{d-1})$. For any $Z \in \mathcal{Z}_{d-1}$, a $(d-1)$ -cycle on $[n]$, it holds that

$$\text{Fill}(Z, T_n^d) = \text{Cone}_n(F) \oplus \text{Fill}(Z', T_{n-1}^d),$$

where $F = \text{Fill}(\text{Link}_n(Z), T_{n-1}^{d-1})$ and $Z' = Z \oplus \text{Cone}_n \text{Link}_n(Z) \oplus F$.

Proof. Observe that since $\text{Link}_n(Z)$ is a $(d-1)$ -cycle on $[n-1]$, F is well defined. Observe also that since F is a subset of T_{n-1}^{d-1} , and T_{n-1}^d is a subset of T_n^d , the right-hand-side of the definition is a subset of T_n^d . Thus, due to acyclicity of T_n^d , it suffices to show that the boundary of the right-hand-side is Z .

By (1) and the definition of the Fill operator,

$$\partial \text{Cone}_n(F) = F \oplus \text{Cone}_n(\partial F) = F \oplus \text{Cone}_n \text{Link}_n(Z),$$

and

$$\partial \left(\text{Fill}(Z', T_{n-1}^d) \right) = Z' = Z \oplus \text{Cone}_n \text{Link}_n(Z) \oplus F.$$

The \oplus of the two terms is indeed Z . □

2 The boundaries of hypertrees

The central result of this section is the following characterization of the boundaries of 2-hypertrees:

Theorem 2.1. *A nontrivial 1-cycle Z on $[n]$ is the boundary of a 2-hypertree on $[n]$ if and only if $|Z| \equiv \binom{n-1}{2} \pmod{2}$.*

Observe that the trivial cycle cannot be the boundary of an acyclic set. The theorem will be proved by first establishing the following much more general (but less precise) result.

Given an acyclic set F of d -simplices on $[n]$, henceforth to be called a d -forest, let $\text{co-rank}(F) = \binom{n-1}{d} - |F|$. In particular, if F is a d -hypertree, then $\text{co-rank}(F) = 0$.

We shall use the convention that $\binom{n}{k} = 0$ for $k < 0$.

Theorem 2.2. *For every $d \geq 1$, and every nontrivial $(d-1)$ -cycle Z_n^{d-1} over $[n]$, there exists a collapsible d -forest F_n^d on $[n]$ such that $\partial F_n^d = Z_n^{d-1}$, and $\text{co-rank}(F_n^d) \leq \binom{n-1}{d-2}$, which is typically much smaller than $\binom{n-1}{d}$. Moreover, there are at least $n - 2d$ different such F_n^d 's.*

Proof. The proof proceeds by a double induction on d and n .

Let us first verify the statement for $d = 1$ by an induction on n . It is, of course, a simple exercise; the reason we spell it out here is that the proof of the general case will have a very similar structure.

For $n = 1$ the statement is vacuous, and for $n = 2$ it is trivial. So we assume $n \geq 3$, and that the statement is true for all $k < n$. Given a nontrivial 0-cycle $Z_n^0 \subseteq [n]$ (i.e., a nonempty set of vertices of even size) as the required boundary, we may assume without loss of generality (by relabelling, if necessary) that $n \in Z_n^0$. Pick any $i \in [n-1]$ such that $Z_n^0 \neq \{i, n\}$, and set $Z_{n-1}^0 = Z_n^0 \oplus \{i, n\}$, a nontrivial 0-cycle on $[n-1]$. Let T_{n-1}^1 be a tree on $[n-1]$ with boundary Z_{n-1}^0 . Its existence is ensured by the induction hypothesis. Attaching n to i in T_{n-1}^1 by a new edge (i, n) , yields the required tree T_n^1 with boundary Z_n^0 . Moreover, the $(n-2)$ possible choices of i yield $(n-2)$ different trees with this property.

Having established the base case of our induction, we proceed to prove the general case (d, n) , $d > 1$, assuming that the statement is true for all n in dimensions smaller than d , and for all (d, k) with $k < n$.

When $d+1 \leq n \leq 2d-1$, it holds that $\binom{n-1}{d} \leq \binom{n-1}{d-2}$, and so the co-rank of any d -forest on $[n]$ is at most $\binom{n-1}{d-2}$. Thus, $F_n^d = \text{Fill}(Z_n^{d-1}, T_n^d)$, where T_n^d is any collapsible d -hypertree, is a collapsible d -forest satisfying the requirement $\partial F_n^d = Z_n^{d-1}$. The number of such forests is obviously $\geq 1 > n - 2d$. Thus, the statement is true for such n 's.

We proceed to the induction step assuming $n \geq 2d$. As before, one may assume without loss of generality (by relabeling the vertices) that the vertex n participates in Z_n^{d-1} . Define $Z_{n-1}^{d-2} = \text{Link}_n(Z_n^{d-1})$; it is a nontrivial $(d-2)$ -cycle on $[n-1]$. Now, consider collapsible $(d-1)$ -forests Y on $[n-1]$ satisfying the three conditions below.

$$\begin{aligned} \text{(a)} \quad \partial(Y) &= Z_{n-1}^{d-2}; \\ \text{(b)} \quad Y &\neq Z_{n-1}^{d-1} \oplus \text{Cone}_n(Z_{n-1}^{d-2}); \\ \text{(c)} \quad |Y| &\geq \binom{n-2}{d-1} - \binom{n-2}{d-3}. \end{aligned} \tag{2}$$

By induction hypothesis on smaller dimensions, there are at least $(n-2d+1)$ different Y 's satisfying (2a) and (2c). Since at most one forest can violate (2b), there are at least $n - 2d$ different Y that satisfy all the three conditions above.

If such Y 's exist, set $F_{n-1}^{d-1} = Y$ for some possible choice of Y . Each such choice will result in a different final construction. Define

$$Z_{n-1}^{d-1} = Z_n^{d-1} \oplus \text{Cone}_n(Z_{n-1}^{d-2}) \oplus F_{n-1}^{d-1}. \tag{3}$$

Observe that (2a) ensures that Z_{n-1}^{d-1} is a $(d-1)$ -cycle on $[n-1]$, and (2b) ensures that it is nontrivial. The induction hypothesis about the case $(d, n-1)$ yields the existence of a collapsible forest F_{n-1}^d such that $\partial F_{n-1}^d = Z_{n-1}^{d-1}$ and $|F_{n-1}^d| \geq \binom{n-2}{d} - \binom{n-2}{d-2}$.

If Y as above does not exist, which may happen only when $n = 2d$, there still exists Y' satisfying (2a) and (2c), and in addition $Y' = Z_n^{d-1} \oplus \text{Cone}_n(Z_{n-1}^{d-2})$. Set $F_{n-1}^{d-1} = Y'$. The cycle Z_{n-1}^{d-1} is still defined by (3), but now it is trivial. Choosing $F_{n-1}^d = \emptyset$, it still holds that $\partial F_{n-1}^d = Z_{n-1}^{d-1}$, and that $|F_{n-1}^d| \geq \binom{n-2}{d} - \binom{n-2}{d-2} = 0$.

Having defined F_{n-1}^{d-1} and F_{n-1}^d (and now it does not matter whether Y exists or not), we finally define the desired F_n^d :

$$F_n^d = \text{Ext}_n(F_{n-1}^d, F_{n-1}^{d-1}) = F_{n-1}^d \oplus \text{Cone}_n(F_{n-1}^{d-1}). \quad (4)$$

Since both F_{n-1}^{d-1} and F_{n-1}^d are collapsible, so is F_n^d .

Next, let us verify that F_n^d has the required boundary and co-rank.

$$\begin{aligned} \partial(F_n^d) &= \partial(F_{n-1}^d) \oplus F_{n-1}^{d-1} \oplus \text{Cone}_n(\partial(F_{n-1}^{d-1})) && \text{(by (1) and the linearity of } \partial) \\ &= Z_{n-1}^{d-1} \oplus F_{n-1}^{d-1} \oplus \text{Cone}_n(Z_{n-1}^{d-2}) && \text{(by definition of } F_{n-1}^d \text{ and (2a))} \\ &= Z_n^{d-1}. && \text{(by (3))} \end{aligned}$$

Further, we have

$$\begin{aligned} \binom{n-1}{d} - |F_n^d| &= \binom{n-1}{d} - (|F_{n-1}^d| + |F_{n-1}^{d-1}|) \\ &= ((\binom{n-2}{d} - |F_{n-1}^d|) + (\binom{n-2}{d-1} - |F_{n-1}^{d-1}|)) \\ &\leq \binom{n-2}{d-2} + \binom{n-2}{d-3} \\ &= \binom{n-1}{d-2}. \end{aligned}$$

Finally, it is easy to verify that each of the $n - 2d$ different choices available for F_{n-1}^{d-1} , yields a different F_n^d , and so there are at least $n - 2d$ different F_n^d with the required properties. \square

It is now an easy matter to prove Theorem 2.1.

Proof. (of Theorem 2.1) By Theorem 2.2 applied to $d = 2$, the size of the constructed F_n^2 with $\partial F_n^2 = Z_n^1$ is either $\binom{n-1}{2}$, or $\binom{n-1}{2} - 1$. However, since the boundary of every 2-simplex consists of 3 edges, $|F_n^2|$ and $|\partial F_n^2|$ must have the same parity. Thus, $|F_n^2| = \binom{n-1}{2}$, i.e., it is a 2-hypertree, if and only if $|Z_n^1|$ has the same parity as $\binom{n-1}{2}$. \square

We conclude this section with the following easy corollary of Theorem 2.2 about the density of boundaries of d -hypertrees among the d -cycles:

Corollary 2.3. *For every $(d-1)$ -cycle Z on $[n]$, there exists a $(d-1)$ -cycle Z' on $[n]$, such that Z' is the boundary of a d -hypertree on $[n]$, and $|Z \oplus Z'| \leq (d+1) \cdot \binom{n-1}{d-2}$.*

3 On the maximum size of a simple d -cycle on $[n]$

Observe that if the boundary of a d -forest F_n^d is of the form $\partial\sigma$ for some d -simplex σ on $[n]$, then $F_n^d \oplus \sigma$ is a simple cycle. Therefore, applying Theorem 2.2 to $(d-1)$ -cycles of the form $\partial\sigma$, we conclude that:

Theorem 3.1.

- Hamiltonian 2-cycles on $[n]$ exist over \mathbb{F}_2 if and only if $n \equiv 0$ or $3 \pmod{4}$.
- For general d , the maximum size of a simple d -cycle on $[n]$ is at least $\binom{n-1}{d} - \binom{n-1}{d-2} + 1$.

Remark. We conclude this section by mentioning that the parity condition $|X| \equiv |\partial(X)| \pmod{2}$ for even dimensions is not the only factor that restricts the size of a simple cycle. The dual of a simple cycle is a cut and when $d = n - 3$, the dual of a simple d -cycle on $[n]$ is a graphical (1-dimensional) cut. Since the largest cut in a graph on n vertices has size at most $n^2/4$, this is the size of a largest simple n -vertex $(n - 3)$ -cycle too. It is shown in [19] that the largest 2-dimensional cut over \mathbb{F}_2 has a size $\binom{n-1}{3} - \frac{n^2}{4} - \Theta(n)$. Hence the size of a largest simple n -vertex $(n - 4)$ -cycle over \mathbb{F}_2 is also $\binom{n-1}{n-4} - \frac{n^2}{4} - \Theta(n)$.

4 Trees with large filling-volume

In this section our aim is to demonstrate that, in every dimension d , there exist trees of average filling-volume $\Omega(n^d)$. We construct a sequence of nice d -trees by conical extensions along carefully chosen relabellings of a $(d - 1)$ -tree of average filling-volume $\Omega(n^{d-1})$. The subtlety lies precisely in the choice and the analysis of a suitable relabelling scheme².

In order to illustrate the main ideas in a more concrete setting, we describe the construction of a family of 2-trees on $[n]$ with average filling-volume $\Omega(n^2)$. In $d = 1$, a Hamiltonian path on $[n]$ has average filling volume $\Omega(n)$. We will show, in particular, how the problem for 2-dimensional trees is reduced to that of constructing a special sequence of Hamiltonian paths.

We use the conical extension described in Section 1.2 multiple times to construct a 2-tree T_n^2 on $[n]$, given a 2-tree T_m^2 on $[m]$, $m < n$, and a sequence of 1-trees T_i^1 on $[i]$, $i \in \{m, \dots, n - 1\}$, as follows:

$$T_{i+1}^2 = \text{Ext}_{i+1}(T_i^2, T_i^1), \quad i \in \{m, \dots, n - 1\}. \quad (5)$$

Specifically, we will assume that n is divisible by 4, to avoid the use of floor and ceil, and choose $m = n/2$, and T_m^2 to be an arbitrary tree on $[m]$.

Our aim is to construct the sequence of 1-dimensional (graphical) trees T_i^1 , to ensure that for a typical 2-simplex $\sigma = \{a, b, c\}$, $\text{Fill}_n^2 = \text{Fill}(\partial\sigma, T_n^2)$ will be large. For that, fixing σ , and using repeatedly Claim 1.2 we get:

$$\text{Fill}_n^2 = \bigoplus_{i=n}^{m+1} \text{Cone}_i(\text{Fill}_{i-1}^1) \oplus \text{Fill}_m^2, \quad (6)$$

where, going by decreasing i , Z_n is the (graphic) triangle $= \{(a, b), (b, c), (c, a)\} = \partial\sigma$, $\text{Fill}_{i-1}^1 = \text{Fill}(\text{Link}_i(Z_i), T_{i-1}^1)$, for $Z_i = Z_{i+1} + \partial \text{Cone}_{i+1}(\text{Fill}_i^1)$, and $\text{Fill}_m^2 = \text{Fill}(Z_m, T_m^2)$. It may be noted that the summands in Equation 6 are a disjoint collection of d -simplices.

Since Fill_m^2 is arbitrary, to obtain a large filling-volume, we rely solely on the sum $\sum_{i=m+1}^n |\text{Cone}_i(\text{Fill}_{i-1}^1)| = \sum_{i=m+1}^n |\text{Fill}_{i-1}^1|$. Hence it is enough to understand the sequence Fill_{i-1}^1 , $i = n, \dots, m + 1$.

Now, Z_i is a 1-dimensional cycle (that is, a graphic cycle). In our case, by construction, we will ensure that it will always be a simple cycle. Hence $\text{Link}_i(Z_i)$ are the two neighbors x, y of i , in Z_i if $i \in Z_i$, and \emptyset if $i \notin Z_i$. In the latter case, we gain nothing to the sum above, hence we will design Z_i so that i will belong to Z_i for many i 's.

Suppose first that $i \in Z_i$, with neighbors x, y in Z_i . Then $\text{Fill}(\{x, y\}, T_{i-1}^1)$ is the path in T_{i-1}^1 from x to y , and in order to contribute significantly to the sum, we wish that this will be large. Obviously for T_{i-1}^1 , being a Hamiltonian path, the “typical” filling-volume will be as large as possible, which motivates the choice of T_{i-1}^1 to be a Hamiltonian path on $[i - 1]$.

Let us examine a concrete construction. Let $A = \{1, \dots, \frac{n}{4}\}$, $B = \{\frac{n}{4} + 1, \dots, \frac{n}{2}\}$ and $C = \{\frac{n}{2} + 1, \dots, n - 1\}$, then each T_i^1 is a path of the form $P_A i P_B P_C$, where P_A , P_B and P_C are Hamiltonian paths on A , B and $C \cap [i - 1]$, respectively (Figure 1). Most importantly, note how P_A and P_B change with the parity of i .

²Let us mention that, for a fixed d , choosing the relabellings independently and uniformly at random in each step $i > d + 1$, yields a sequence $\{T_i^d\}_{i=d+2}^\infty$ of random d -trees, where every T_i^d has the desired property with positive probability. Interestingly, more can be said in this setting. Let Z be any $(d - 1)$ -cycle in K_n^{d-1} , and let \tilde{T}_n^d be a random uniform relabeling of T_n^d from this sequence. Then, the expected value of $\text{Fill}(Z, \tilde{T}_n^d)$ is $\Omega(n^d)$. The analysis is omitted here in favor of the explicit deterministic construction presented next.

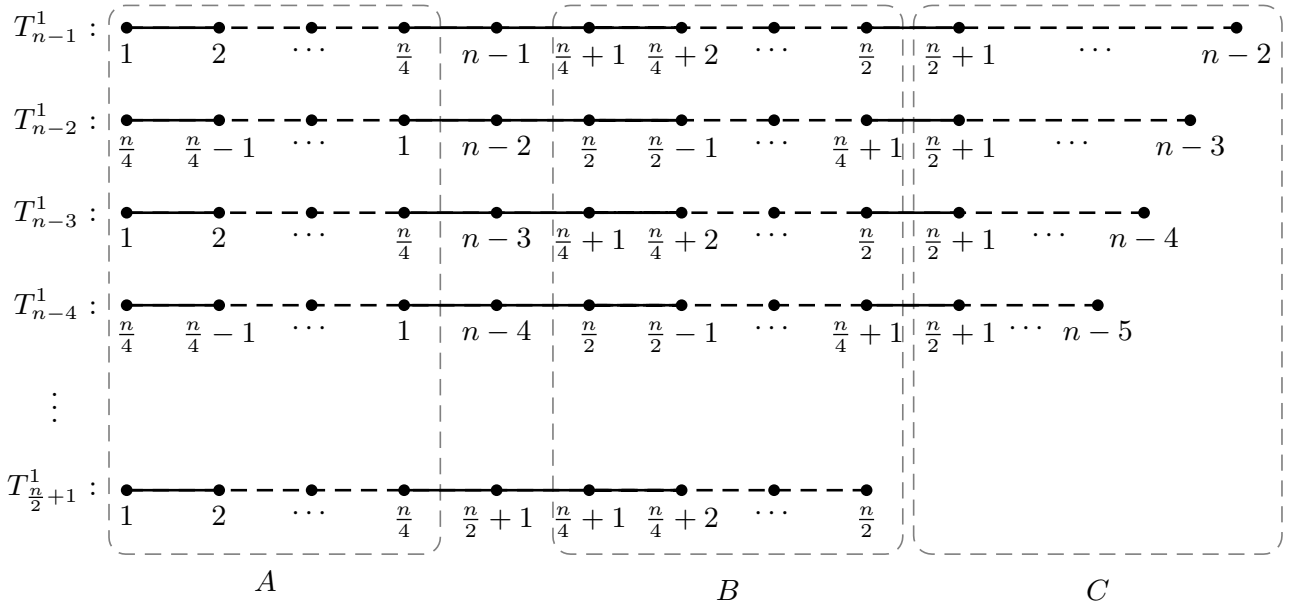


Figure 1: The sequence of 1-trees $\{T_i^1\}$ used to construct a 2-tree on n vertices with average filling-volume $\Omega(n^2)$.

We call a 2-simplex $\sigma = \{a, b, c\}$ “good” if $a \in A$, $b \in B$ and $c \geq n/2 + 2$. Note that for a good $\sigma = \{a, b, c\}$, if $c < n$, $c \notin Z_n = \partial\sigma$ and hence $Z_{n-1} = Z_n = \sigma$. Thus this carries on as long as $i > c$. For $i = c$, $\text{Link}_i(Z_i) = \{a, b\}$, and then $Z_{i-1} = \{(a, b)\} \cup \text{Fill}(\{a, b\}, T_i^1)$ which is, by construction, the simple cycle containing the suffix of P_A from a to $i-1$, then the prefix of P_B from $i-1$ to b , plus the edge (a, b) . In particular, it is simple and contains $i-1$. Note also, that by the alteration between even and odd values of i , from $i = c$ down to $n/2 + 1$, $\text{Link}_j(Z_j)$ are $\{1, \frac{n}{2}\}$ for even j , and $\{\frac{n}{4}, \frac{n}{4} + 1\}$ for odd j , for every $\frac{n}{2} + 1 < j < c$. This is summed up in the following:

Claim 4.1.

$$\text{Link}_i(Z_i^1) = \begin{cases} \emptyset, & i \in \{c+1, \dots, n\}, \\ \{a, b\}, & i = c, \\ \{\frac{n}{4}, \frac{n}{4} + 1\}, & i \in \{\frac{n}{2} + 1, \dots, c-1\}, i \text{ odd, and} \\ \{1, \frac{n}{2}\}, & i \in \{\frac{n}{2} + 1, \dots, c-1\}, i \text{ even.} \end{cases} \quad (7)$$

Moreover,

$$|\text{Fill}_i^1| = \frac{n}{2} + 1, \quad i \in \left\{\frac{n}{2} + 1, \dots, c-1\right\}. \quad (8)$$

Hence, equation 6 implies that for a good σ , $|\text{Fill}(\{a, b, c\}, T_n^2)| \gtrsim (c - \frac{n}{2} - 2) \frac{n}{2}$.

For any $c \geq n/2 + 2$, the number of good triangles containing c is $|A||B| = (\frac{n}{4})(\frac{n}{4})$. Hence

$$\begin{aligned} \sum_{\sigma \text{ good}} |\text{Fill}(\sigma, T_n^2)| &\gtrsim \sum_{c=n/2+3}^n \frac{n^2}{16} (c - \frac{n}{2} - 2) \frac{n}{2} \\ &\gtrsim n^5/2^9. \end{aligned} \quad (9)$$

Since the total number of 2-faces is less than n^3 , we see that $\mu(T_n^2) \in \Omega(n^2)$.

After introducing one more notation, let us briefly discuss the essential features in the construction of T_n^2 that resulted in a large average filling-volume. Given a tree T in K_n^d and a d -simplex $\tau \in T$, the set of all d -simplices in K_n^d such that $T \oplus \tau \oplus \sigma$ is a tree, is called the *cut of τ in T* , denoted by $\text{Cut}(\tau, T)$.

We constructed T_n^2 by repeated conical extensions of an arbitrary 2-tree $T_{n/2+1}^2$ on $[\frac{n}{2} + 1]$ using the sequence of 1-trees depicted in Figure 1. It should be clear that we only used the properties of this sequence of 1-trees to demonstrate that the average filling-volume of T_n^2 is $\Omega(n^2)$. This will be our approach in higher dimensions too. We start with an arbitrary d -tree on $[m]$ where this time

$m = \lceil (1 - 1/d)n \rceil$ and then extend it to a d -tree on $[n]$ along a sequence of carefully chosen $(d-1)$ -trees T_i^{d-1} , $i \in \{m, \dots, n-1\}$.

If we look at the initial $\frac{n}{2}$ -length segment of each of the 1-trees in Figure 1, we see that for every odd i , this initial segment consists of a path $X = (1, \dots, \frac{n}{2})$ with the center edge $x = \{\frac{n}{4}, \frac{n}{4} + 1\}$ replaced by the 2-length path $\text{Cone}_i(\partial x)$. Similarly, for even i , the initial segment consists of a path $Y = (\frac{n}{4}, \dots, 1, \frac{n}{2}, \dots, \frac{n}{4} + 1)$ with the center edge $y = \{1, \frac{n}{2}\}$ replaced by the 2-length path $\text{Cone}_i(\partial y)$. Let us treat X and Y , rightfully, as 1-trees over $[n/2]$ and note the following three useful properties. Firstly, $\text{Cut}(x, X)$ and $\text{Cut}(y, Y)$ are very large ($n^2/16$, namely of order of the total number of 1-simplices) and this helps in having a constant fraction of 2-faces which are “good” for the final 2-tree. Secondly, observe that $\text{Fill}(y, X)$ and $\text{Fill}(x, Y)$ are large ($n/2$; of order of a 1-tree on n vertices) and hence contribute a large number of simplices to the filling of a good 2-simplex at each level. Finally observe that $\text{Fill}(y, X)$ contains x and $\text{Fill}(x, Y)$ contains y so that $\text{Fill}(z, T_i^1)$ contains $i-1$ where $z = y$ when i is odd and $z = x$ otherwise, for every $i \in \{\frac{n}{2} + 1, \dots, n-1\}$. Notice also, that the explicit structure of T_i^1 on the vertices in C was not used, and may be replaced with an arbitrary maximal acyclic extension. We also note that while the fact that Z_i was simple facilitated an explicit description of Fill_i^1 , this is not really an essential part of the proof, once one observes the properties of the edges x and y in the odd and even trees respectively.

For $d \geq 3$, it is difficult to explicitly describe a pair of $(d-1)$ -trees with the above properties, but we show that from any $(d-1)$ -tree with a large average filling-volume, we can construct a pair of $(d-1)$ -trees which have the above properties. Firstly, we show that in any tree with a large filling-volume, there is a simplex which is part of a large number of fillings among which at least one is large. A carefully chosen isomorphism of such a tree will give its pair. Notice that given a tree T and $\tau \in T$, the number of fillings in which τ participates is $|\text{Cut}(\tau, T)|$.

Lemma 4.2. *In any d -tree T on n vertices with $\mu(T) \geq c \binom{n-1}{d}$, there exist $\tau \in T$ and $\sigma \in \text{Cut}(\tau, T)$ such that $|\text{Cut}(\tau, T)| \cdot |\text{Fill}(\sigma, T)| \geq \frac{c^3}{8} \binom{n}{d+1} \binom{n-1}{d}$.*

Proof. For any $\tau \in T$, let $f(\tau, T) = \sum_{\sigma \in \text{Cut}(\tau, T)} |\text{Fill}(\sigma, T)|$. Let σ^* be a d -simplex which produces a largest filling in T among all the simplices in $\text{Cut}(\tau, T)$. Since $|\text{Cut}(\tau, T)| \cdot |\text{Fill}(\sigma^*, T)| \geq f(\tau, T)$, by the remark above, it suffices to show that there exists $\tau \in T$ with $f(\tau, T) \geq \frac{c^3}{8} \binom{n}{d+1} \binom{n-1}{d}$.

If we sum $f(\tau, T)$ over all $\tau \in T$, we add the sizes of fillings of T , with each filling being added once for each d -face contained in it. That is $\sum_{\tau \in T} f(\tau, T) = \sum_{\sigma \in K_n^d} |\text{Fill}(\sigma, T)|^2$, which is the square of the l_2 -norm of the filling-volumes. Since the maximum size of a filling in T is $\binom{n-1}{d}$ (the size of T) and the average filling-volume of T is at least $c \binom{n-1}{d}$ (by assumption), it is not difficult to see that at least $c/2$ fraction of the fillings have a size at least half the average. Hence $\sum_{\tau \in T} f(\tau, T)$ is at least $\left(\frac{c}{2} \binom{n}{d+1}\right) \left(\frac{c}{2} \binom{n-1}{d}\right)^2 = \frac{c^3}{8} \binom{n}{d+1} \binom{n-1}{d}^2$. By averaging there exists at least one τ in T which satisfies our required bound. \square

By a systematic extension of the 2-dimensional construction, keeping intact the essential features noted above, we establish the following theorem; a complete proof of which is given in the appendix.

Theorem 4.3. *For any two positive integers d and n , there exists a d -tree T on $[n]$ with $\mu(T) \geq c_d \binom{n-1}{d}$, where $c_d = 16(48)^{-3^{d-1}}$.*

5 A remark on cycles and hypertrees over \mathbb{Q} .

We conclude by pointing out that our results over \mathbb{F}_2 extend to \mathbb{Q} . Although cycles over \mathbb{F}_2 are not necessarily cycles over \mathbb{Q} , one can verify that if Z is a simple \mathbb{F}_2 -cycle, then it is contained in the support of a simple \mathbb{Q} -cycle. This follows from the \mathbb{F}_2 -acyclicity of $Z' = Z \setminus \{\sigma\}$ for any simplex $\sigma \in Z$, which in turn implies \mathbb{Q} -acyclicity for Z' . Hence (i) a Hamiltonian d -cycle over \mathbb{Q} exists whenever a Hamiltonian d -cycle over \mathbb{F}_2 exists and (ii) the average filling-volume of a d -tree T computed over \mathbb{Q} is at least as big as the average filling-volume of T computed over \mathbb{F}_2 .

It is possible that simple cycles over \mathbb{Q} may be larger than largest simple cycles over \mathbb{F}_2 . We know that this is indeed the case when $d = n - 4$ from a work on hypercuts, which are duals of simple

cycles [19]. We believe that for $d = 2$, we can have an n -vertex Hamiltonian cycle over \mathbb{Q} for all large n , in contrast with the \mathbb{F}_2 -case. Though the construction in the proof of Theorem 2.2 is done over \mathbb{F}_2 , with a proper generalization of the cone and link operators, we can do this construction over any field. Let us call a chain over \mathbb{Q} whose support is a \mathbb{Q} -tree a \mathbb{Q} -weighted tree. There are no parity restrictions over \mathbb{Q} , but one can show that when $n \leq 5$, all the boundaries forbidden for n -vertex 2-trees over \mathbb{F}_2 are forbidden for n -vertex \mathbb{Q} -weighted 2-trees too. One can establish that these are the only restrictions over \mathbb{Q} . That is,

Claim 5.1. *Given any 1-cycle Z over \mathbb{Q} on $[n]$, $n \geq 6$, we can construct a \mathbb{Q} -weighted tree T on $[n]$ with $\partial T = Z$. In particular, n -vertex Hamiltonian 2-cycles over \mathbb{Q} exist for every $n \geq 6$.*

The proof idea is to use a case analysis to establish the claim for $n = 6$ and then use the conical extension to construct a \mathbb{Q} -weighted tree whose boundary is Z . This will have an immediate effect in higher dimensions on the co-rank term in a result for \mathbb{Q} analogous to Theorem 2.2.

References

- [1] Lior Aronshtam and Nathan Linial. The threshold for d -collapsibility in random complexes. *Random Structures & Algorithms*, 2015.
- [2] Lior Aronshtam and Nathan Linial. When does the top homology of a random simplicial complex vanish? *Random Structures & Algorithms*, 46(1):26–35, 2015.
- [3] Lior Aronshtam, Nathan Linial, Tomasz Łuczak, and Roy Meshulam. Collapsibility and vanishing of top homology in random simplicial complexes. *Discrete & Computational Geometry*, 49(2):317–334, 2013.
- [4] Eric Babson, Christopher Hoffman, and Matthew Kahle. The fundamental group of random 2-complexes. *Journal of the American Mathematical Society*, 24(1):1–28, 2011.
- [5] Anders Björner. Topological methods. *Handbook of combinatorics*, vol. 1, 2, 1819–1872, 1995.
- [6] Martin Čadek, Marek Krčál, Jiří Matoušek, Lukáš Vokřínek, and Uli Wagner. Extending continuous maps: Polynomiality and undecidability. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 595–604. ACM, 2013.
- [7] Gunnar Carlsson. Topology and data. *Bulletin of the American Mathematical Society*, 46(2):255–308, 2009.
- [8] Tamal K Dey. *Curve and surface reconstruction: Algorithms with mathematical analysis*, volume 23. Cambridge University Press, 2006.
- [9] Jacob Fox, Mikhail Gromov, Vincent Lafforgue, and Assaf Naor. Overlap properties of geometric expanders. In *Proc. 22nd ACM-SIAM Symposium on Discrete Algorithms*, number EPFL-CONF-170579, pages 1188–1197. Walter De Gruyter & Co, 2011.
- [10] Robert Ghrist. Elementary applied topology. *Book in preperation*, 2014.
- [11] Mikhail Gromov. Singularities, expanders and topology of maps. Part 2: From combinatorics to topology via algebraic isoperimetry. *Geometric and Functional Analysis*, 20(2):416–526, 2010.
- [12] Maurice Herlihy and Nir Shavit. The topological structure of asynchronous computability. *Journal of the ACM (JACM)*, 46(6):858–923, 1999.
- [13] Jeff Kahn, Michael Saks, and Dean Sturtevant. A topological approach to evasiveness. *Combinatorica*, 4(4):297–306, 1984.
- [14] Gil Kalai. Enumeration of Q -acyclic simplicial complexes. *Israel Journal of Mathematics*, 45(4):337–351, 1983.

- [15] Tali Kaufman, David Kazhdan, and Alexander Lubotzky. Ramanujan complexes and bounded degree topological expanders. In *Foundations of Computer Science (FOCS), 2014 IEEE 55th Annual Symposium on*, pages 484–493. IEEE, 2014.
- [16] Nathan Linial and Roy Meshulam. Homological connectivity of random 2-complexes. *Combinatorica*, 26(4):475–487, 2006.
- [17] Nathan Linial, Roy Meshulam, and Mishael Rosenthal. Sum complexes - a new family of hypertrees. *Discrete & Computational Geometry*, 44(3):622–636, 2010.
- [18] Nathan Linial and Yuval Peled. On the phase transition in random simplicial complexes. *arXiv preprint arXiv:1410.1281*, 2014.
- [19] Nati Linial, Ilan Newman, Yuval Peled, and Yuri Rabinovich. Extremal problems on shadows and hypercuts in simplicial complexes. *arXiv preprint arXiv:1408.0602*, 2014.
- [20] Alexander Lubotzky and Roy Meshulam. Random Latin squares and 2-dimensional expanders. *Advances in Mathematics*, 272:743–760, 2015.
- [21] Jiří Matoušek, Eric Sedgwick, Martin Tancer, and Uli Wagner. Embeddability in the 3-sphere is decidable. In *Proceedings of the thirtieth annual symposium on Computational geometry*, page 78. ACM, 2014.
- [22] Jiří Matoušek, Martin Tancer, and Uli Wagner. Hardness of embedding simplicial complexes in \mathbb{R}^d . *Journal of the European Mathematical Society*, 13(2):259–295, 2011.
- [23] Jiří Matoušek. *Using the Borsuk-Ulam theorem: Lectures on topological methods in combinatorics and geometry*. Springer Science & Business Media, 2008.
- [24] Roy Meshulam and Nathan Wallach. Homological connectivity of random k -dimensional complexes. *Random Structures & Algorithms*, 34(3):408–417, 2009.
- [25] Ilan Newman and Yuri Rabinovich. On multiplicative λ -approximations and some geometric applications. *SIAM Journal on Computing*, 42(3):855–883, 2013.
- [26] Ilan I Newman and Yuri Rabinovich. On connectivity of the facet graphs of simplicial complexes. *arXiv preprint arXiv:1502.02232*, 2015.
- [27] Michael Saks and Fotios Zaharoglou. Wait-free k -set agreement is impossible: The topology of public knowledge. *SIAM Journal on Computing*, 29(5):1449–1483, 2000.

A Proof of Theorem 4.3

Statement. For any two positive integers d and n , there exists a d -tree T on $[n]$ with $\mu(T) \geq c_d \binom{n-1}{d}$, where $c_d = 16(48)^{-3^{d-1}}$.

Proof. For $d = 1$, the required tree T is the Hamiltonian path on n vertices. It is easy to verify that $\mu(T) \geq \frac{1}{3}(n-1)$ as claimed. Hence let $d \geq 2$ and we assume that the statement of the theorem is true in dimension $d-1$. In particular, we will use a $(d-1)$ -tree X on m vertices, $m = \lceil (1-1/d)n \rceil$, with large average filling-volume in the construction of T . Since $c_d \binom{n-1}{d} \leq 1$ for $n < 10d$, we assume $n \geq 10d$. We set $c = c_{d-1}$ (for readability).

We construct a sequence of d -trees, $T_{m+1}^d, T_{m+2}^d, \dots, T_n^d = T$, where T_{m+1}^d is an arbitrary d -tree on $[m+1]$ and each T_{i+1}^d , $i \in \{m+1, \dots, n-1\}$, is a conical extension of T_i^d . That is,

$$T_{i+1}^d = \text{Ext}_{i+1}(T_i^d, T_i^{d-1}), \quad i \in \{m+1, \dots, n-1\}, \quad (10)$$

where T_i^{d-1} is a $(d-1)$ -tree on $[i]$ which we construct as described next.

For each $i \geq m+1$, we construct T_i^{d-1} based on a $(d-1)$ -tree X on the vertex set $[m]$ which has a large average filling-volume, i.e., $\mu(X) \geq c \binom{m-1}{d-1}$. Such a tree X exists by the induction hypothesis. By Lemma 4.2, there exist $x \in X$ and $y \in \text{Cut}(x, X)$ such that $|\text{Cut}(x, X)| \cdot |\text{Fill}(y, X)| \geq \frac{c^3}{8} \binom{m-1}{d-1} \binom{m}{d}$. Let Y be a $(d-1)$ -tree isomorphic to X obtained from X by relabelling its vertices so that x and y exchange their roles. That is, $y \in Y$ and $x \in \text{Cut}(y, Y)$ with $|\text{Cut}(y, Y)| \cdot |\text{Fill}(x, Y)| \geq \frac{c^3}{8} \binom{m-1}{d-1} \binom{m}{d}$. For an odd number i in $\{m+1, \dots, n-1\}$, we construct a forest F_i on the vertex set $[m] \cup \{i\}$ by removing x from X and adding $\text{Cone}_i(\partial x)$ in its place, i.e., $F_i = X \oplus x \oplus \text{Cone}_i(\partial x)$. Note that since X is acyclic, F_i is acyclic. Though it is not a tree on $[m] \cup \{i\}$ yet, any $(d-1)$ -face $\sigma \in \text{Cut}(x, X)$ creates a cycle in F_i . More precisely, for any $\sigma \in \text{Cut}(x, X)$,

$$\text{Fill}(\sigma, F_i) = \text{Fill}(\sigma, X) \oplus x \oplus \text{Cone}_i(\partial x). \quad (11)$$

Similarly, for an even number i in $\{m+1, \dots, n-1\}$, we construct a forest F_i on the vertex set $[m] \cup \{i\}$ by removing y from Y and adding $\text{Cone}_i(\partial y)$ in its place, i.e., $F_i = Y \oplus y \oplus \text{Cone}_i(\partial y)$. Note that F_i is acyclic in this case too and for any $\sigma \in \text{Cut}(y, Y)$,

$$\text{Fill}(\sigma, F_i) = \text{Fill}(\sigma, Y) \oplus y \oplus \text{Cone}_i(\partial y), \quad (12)$$

Finally, we let T_i^{d-1} be any $(d-1)$ -tree on $[i]$ containing F_i . This completes the construction of T_n^d and we are all set to estimate $\mu(T_n^d)$. We note here that this construction is a generalization of that in the explicit proof for $d = 2$. The forests F_i are just the trees T_i with out the extension to the part in C and the faces x and y are just edges $\{\frac{n}{4}, \frac{n}{4} + 1\}$ and $\{1, \frac{n}{2}\}$ respectively.

Claim A.0.1. Let Γ be the set of (“good”) d -faces of the form $\{i\} \cup \gamma'$ where $i \in \{m+3, \dots, n\}$ and γ' belongs to $\text{Cut}(x, X)$ if i is even and $\text{Cut}(y, Y)$ if i is odd. Then, for each $\gamma = \{i\} \cup \gamma' \in \Gamma$,

$$|\text{Fill}(\gamma, T_n^d)| \geq |\text{Fill}(y, X)| \cdot (i - m - 2).$$

We prove the above claim by establishing that $\text{Fill}(\gamma, T_n^d)$ contains the following “large” set K_γ defined below.

$$K_\gamma = \bigoplus_{j=m+1}^{i-1} \text{Cone}_{j+1}(\text{Fill}(\sigma_j, F_j)), \quad (13)$$

where $\sigma_{i-1} = \gamma'$ and for all $j < i-1$, σ_j is y when j is odd and x otherwise. Notice that $|K_\gamma| \geq |\text{Fill}(y, X)| \cdot (i - m - 2)$ and hence we will establish Claim A.0.1 if we show that K_γ is contained in $\text{Fill}(\gamma, T_n^d)$. It follows from Equations (11) and (12) that for $j \geq m+1$,

$$\begin{aligned} \partial \text{Cone}_{j+1}(\text{Fill}(\sigma_j, F_j)) &= \text{Cone}_{j+1}(\partial \sigma_j) \oplus \text{Fill}(\sigma_j, F_j) \\ &= \text{Cone}_{j+1}(\partial \sigma_j) \oplus H_j \oplus \text{Cone}_j(\partial \sigma_{j-1}) \end{aligned} \quad (14)$$

where H_j is some set of $(d-1)$ -simplices on $[m]$. Substituting Equation (14) in Equation (13) results in a partial telescoping giving rise to the following.

$$\begin{aligned}\partial(K_\gamma) &= \text{Cone}_i(\partial\gamma') \oplus \text{Cone}_{m+1}(\partial\sigma_m) \oplus \sum_{j=m+1}^{i-1} H_j \\ &= \partial\gamma \oplus \gamma' \oplus \text{Cone}_{m+1}(\partial\sigma_m) \oplus \sum_{j=m+1}^{i-1} H_j \\ &= \partial\gamma \oplus Z\end{aligned}$$

where Z is a collection of $(d-1)$ -simplices on $[m+1]$. Since $Z = \partial(K_\gamma) + \partial\gamma$, being a boundary, is a d -cycle on $[m+1]$ and T_{m+1}^d is a tree on $[m+1]$, there exists a unique filling K' of Z in T_{m+1}^d . Then, since K_γ and K' are disjoint, $\partial(K_\gamma \cup K') = \partial\gamma$. This means, by definition, that $K_\gamma \cup K' = \text{Fill}(\gamma, T_n^d)$ and it completes the proof of Claim A.0.1.

We obtain the required lower bound on the sum of filling-volumes by adding the filling-volumes of all fillings generated by d -faces in Γ . The details of estimation are below.

$$\begin{aligned}\binom{n}{d+1}\mu(T_n^d) &\geq \sum_{\gamma \in \Gamma} |\text{Fill}(\gamma, T_n^d)| \\ &\geq \sum_{i=m+3}^n |\text{Cut}(x, X)| \cdot |\text{Fill}(y, X)|(i-m-2) \quad (\text{Claim A.0.1}) \\ &= |\text{Cut}(x, X)| \cdot |\text{Fill}(y, X)| \cdot \frac{1}{2} \left(\frac{n}{d} - 3\right) \left(\frac{n}{d} - 2\right) \quad (n-m \geq \frac{n}{d} - 1) \\ &\geq \frac{c^3}{2^4} \binom{m}{d} \binom{m-1}{d-1} \frac{(n-3d)(n-2d)}{d^2} \quad (\text{By choice of } x \text{ and } y) \\ &\geq \frac{c^3}{2^4} \left(1 - \frac{1}{d}\right)^{2d-1} \binom{n-1}{d} \binom{n-2}{d-1} \frac{(n-3d)(n-2d)}{d^2} \quad (\text{Observation A.1}) \\ &\geq \frac{c^3}{2^4} \left(1 - \frac{1}{d}\right)^{2d-1} \binom{n-1}{d} \frac{1}{4} \binom{n}{d+1} \quad (n \geq 10d) \\ &\geq \frac{c^3}{2^8} \binom{n-1}{d} \binom{n}{d+1} \quad \left(\left(1 - \frac{1}{d}\right)^{2d-1} \geq \frac{1}{8}, \forall d \geq 2\right).\end{aligned}$$

Notice that $\frac{1}{2^8}(c_{d-1})^3 = c_d$. □

Observation A.1. Let d and n be any two positive integers. Then for $\alpha = (1 - 1/d)$ and $m = \lceil \alpha n \rceil$

$$\begin{aligned}\binom{m}{d} &\geq \alpha^d \binom{n-1}{d}, ; \\ \binom{m-1}{d-1} &\geq \alpha^{d-1} \binom{n-2}{d-1}.\end{aligned}$$

Proof. Both inequalities above follow by noting that $(\alpha n - i) = \alpha(n - i/\alpha) \geq \alpha(n - i - 1)$ for all $i \in \{0, \dots, d-1\}$ and then expanding the binomial coefficient on the left side. □